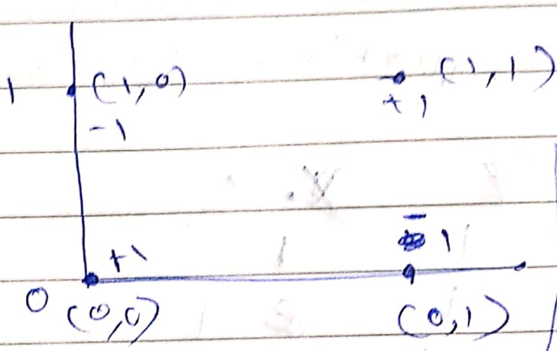
Q3) XOR

A	B	Output
1	1	1
0	1	0
1	0	0
0	0	1

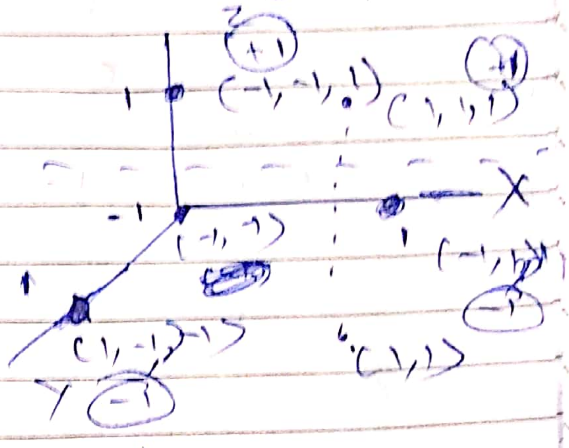


A	B	Output
1	1	1
-1	1	-1
1	-1	-1
-1	-1	1

Assuming 0 = -1 class.
Clearly acc to figure classes are not linearly separable.
=> A kernel is needed.

If we use a new attribute $Z = X \cdot Y$ then the equation becomes linearly separable.

X	Y	X.Y
1	1	1
-1	1	-1
1	-1	-1
-1	-1	1



$\Rightarrow \phi$ Map: $(x, y) \rightarrow (x, y, xy)$

$\Rightarrow$ kernel $K(x_i, x_j) = [x_i \; y_i \; x_i y_i] \begin{bmatrix} x_j \\ y_j \\ x_j y_j \end{bmatrix}$

$= x_i x_j + y_i y_j + x_i x_j y_i y_j$

Gram Matrix

$$K = \begin{array}{c|cccc} & (1,1) & (-1,1) & (1,-1) & (-1,-1) \\ \hline (1,1) & 3 & -1 & -1 & -1 \\ (-1,1) & -1 & 3 & -1 & -1 \\ (1,-1) & -1 & -1 & 3 & -1 \\ (-1,-1) & -1 & -1 & -1 & 3 \end{array}$$

1 = (1,1)
2 = (-1, 1)
3 = (1, -1)
4 = (-1, -1)

$$L_{d}(\alpha) = \sum_{i=1}^{4} \alpha_i - \frac{1}{2} \sum_{i=1}^{4} \sum_{j=1}^{4} \alpha_i \alpha_j y_i y_j K(x_i, x_j)$$

$$= \alpha_1 + \alpha_2 + \alpha_3 + \alpha_4 - \frac{1}{2} \big(3\alpha_1^2 + \alpha_1\alpha_2 + \alpha_1\alpha_3 - \alpha_1\alpha_4 + 3\alpha_2^2 + \alpha_1\alpha_2 + \alpha_2\alpha_4 - \alpha_2\alpha_3 + 3\alpha_3^2 + \alpha_3\alpha_1 + \alpha_3\alpha_4 - \alpha_3\alpha_2 + 3\alpha_4^2 + \alpha_4\alpha_2 + \alpha_4\alpha_3 - \alpha_4\alpha_1 \big)$$

~~$\Rightarrow \alpha_1(3\alpha_1 + \alpha_2 + \alpha_3 - \alpha_4+1) =$~~

$\Rightarrow$ on diff $Loss(\alpha) = \frac{d}{d\alpha}\left(\alpha - \frac{1}{2} H(\alpha^T\alpha)\right)$

$\Rightarrow$ where H = Gram matrix $\times (y_i y_j)$

$\Rightarrow 1 - H\alpha = 0$

$\Rightarrow H\alpha = 1$

$\Rightarrow 3\alpha_1 + \alpha_2 + \alpha_3 - \alpha_4 = 1$

$\alpha_1 + 3\alpha_2 + \alpha_3 + \alpha_4 = 1$

$\alpha_1 + \alpha_2 + 3\alpha_3 + \alpha_4 = 1$

$-\alpha_1 + \alpha_2 + \alpha_3 + 3\alpha_4 = 1$

$\Rightarrow$ optimal $\alpha_1 = \alpha_2 = \alpha_3 = \alpha_4 = \frac{1}{4}$

$\Rightarrow$ All are support vectors, Sub values in Loss equation we get

$Loss(\alpha) = 1 - \frac{1}{2}\left(\frac{3}{16}\times 4 + \frac{1}{16}\times 4\right)$

$= 1 - \frac{1}{2}\left(4 \times \frac{4}{16}\right)$

$= 1 - \frac{1}{2} = \frac{1}{2}$ ~~$= \frac{1}{2}\|w\|^2$~~

$\Rightarrow \vec{w} = \sum_{i=1}^{4} \alpha_i y_i \phi(x_i)$

$= \frac{1}{4}\big((1,1,1) - (-1,1,-1) - (1,-1,-1) + (-1,-1,1)\big)$

$= \frac{1}{4}\big((0,0,4)\big) = (0,0,1)$

$\Rightarrow$ boundary $= \vec{w} \cdot \phi(x) = (0,0,1)\cdot(x, y, x\cdot y)$

$= x \cdot y = 0$

if $x=y=1$ and $x=y=-1$
$\Rightarrow x.y = 1 \times 1 > 0$ & $-1 \times -1 > 0$
$\Rightarrow$ class $= 1$
if $x = 1, y = -1$ & $x = -1, y = 1$

$\Rightarrow x.y = -1 \times 1 < 0$ & $1 \times -1 < 0$
done $\Rightarrow$ class $= -1$

hence XOR done by SVM

Q2) $k(x_i, x_j) = (1 + x^T . x')^2$

where $x^T = [x_1 \; x_2]$
& $x' = \begin{bmatrix} x_1' \\ x_2' \end{bmatrix}$

$\Rightarrow x^T . x' = [x_1 \; x_2] \begin{bmatrix} x_1' \\ x_2' \end{bmatrix} = x_1 x_1' + x_2 x_2'$

$\Rightarrow (1 + x_1 x_1' + x_2 x_2')^2$
$= (1.1) + (x_1^2)(x_1'^2) + (x_2^2)(x_2'^2) + 2(x_1)(x_1') + 2(x_2)(x_2')$
$+ 2(x_1 x_1')(x_2 x_2')$

$\phi(x) . \phi(x') = k(x_i, x_j)$

As we have 6 terms in the expansion of kernel function & each term can be expressed as the product of 2 terms.

$1 = 1.1$, $x_1^2 x_1'^2 = x_1^2 \times x_1'^2$,
$(x_2^2 . x_2'^2) = x_2^2 \times x_2'^2$,
$2 x_1 x_1' = \sqrt{2} x_1 \times \sqrt{2} x_1'$

$2x_2x_2' = \sqrt{2}x_2 \times \sqrt{2}x_2'$

$2x_1x_1'x_2x_2' = \sqrt{2}x_1x_2 \times \sqrt{2}x_1'x_2'$

where each term is a product of one term in x & one term in x'.

$\Rightarrow k(x, x') = [1, x_1^2, x_2^2, \sqrt{2}x_1, \sqrt{2}x_2, \sqrt{2}x_1x_2]^T$
$[1, x_1'^2, x_2'^2, \sqrt{2}x_1', \sqrt{2}x_2', \sqrt{2}x_1'x_2']$

~~$\Rightarrow \phi(x) = [1, x_1^2, \ldots, x_2^2, \sqrt{2}x_1 \ldots$~~

$\Rightarrow \phi(x) = [1, x_1^2, x_2^2, \sqrt{2}x_1, \sqrt{2}x_2, \sqrt{2}x_1x_2]^T$

There are 6 features in $\phi(x)$

Q3) $w = \sum_{i=1}^{m} \alpha_i y_i x_i = \sum_{i=1}^{10} \alpha_i y_i x_i$

$= 0.414 \times 1 \times (4, 2.9) + 0 + 0 + (0.018 \times -1 \times (2.5, 1))$
$+ 0 + 0 + (0.018 \times 1 \times (3.5, 4)) + 0$
$+ (0.414 \times -1 \times (2, 2.1)) + 0$

$= \cancel{\begin{bmatrix} 4.2 \\ 9 \end{bmatrix}} \begin{bmatrix} 1.656 \\ 1.2006 \end{bmatrix} + \begin{bmatrix} -2.45 \\ -1.18 \end{bmatrix} + \begin{bmatrix} 0.063 \\ 0.072 \end{bmatrix}$
$+ \begin{bmatrix} -0.828 \\ -0.8694 \end{bmatrix} = \begin{bmatrix} -2.059 \\ -0.7768 \end{bmatrix}$

2) Support vectors $\Rightarrow \alpha_i \neq 0$
$\Rightarrow x_i = (4, 2.9), (2.5, 1), (3.5, 4), (2, 2.1)$ are all support vectors.

Scanned by CamScanner

CLASSTEAM Page No.
Date / /

Q5) $W = \sum_{i=1}^{n} \alpha_i y_i x_i = \sum_{i=1}^{10} \alpha_i y_i x_i$

$= 0.414 \times (4, 2.9) + 0 + 0$
$+ (1.18 \times -1 \times (2.5, 1)) + 0 + 0$
$+ (1.18 \times 1 \times (3.5, 4)) + 0 +$
$(0.414 \times -1 \times (2, 2.1)) + 0$

$= \begin{bmatrix} 1.656 \\ 1.2006 \end{bmatrix} + \begin{bmatrix} -2.95 \\ -1.18 \end{bmatrix} + \begin{bmatrix} 4.13 \\ 4.72 \end{bmatrix} + \begin{bmatrix} -0.828 \\ -0.869 \end{bmatrix}$

$w = \begin{bmatrix} 2.008 \\ 3.876 \end{bmatrix}$

Support vectors $\Rightarrow \alpha_i \neq 0$
$\Rightarrow x_i = (4, 2.9), (2.5, 1), (3.5, 4), (2, 2.1)$ are all support vectors.

<u>Using KKT cond</u>

$b_i = (1 - W^T x^{(s)}) \times y_i$

$b_1 = 1 - W^T \begin{bmatrix} 4 \\ 2.9 \end{bmatrix} = -18.2724$

$b_2 = -1 - W^T \begin{bmatrix} 2.5 \\ 1 \end{bmatrix} = -9.896$

$b_3 = 1 - W^T \begin{bmatrix} 3.5 \\ 4 \end{bmatrix} = -21.508$

$b_4 = -1 - W^T \begin{bmatrix} 2 \\ 2.1 \end{bmatrix} = -13.1556$

$\Rightarrow$ Avg $= -15.708$

Scanned by CamScanner

for (3,3)

$w^T x + b = 17.652 - 15.708$

$= 1.944 > 0$

$\Rightarrow$ belongs to class 1

Scanned by CamScanner

Q6) BONUS

RBF kernel $= e^{-\gamma \|U - V\|^2}$

$\|U - V\| = \sqrt{(U_1 - V_1)^2 + (U_2 - V_2)^2 \ldots}$

but U & V are 1-D

$\Rightarrow \|U - V\| = \sqrt{(U - V)^2} = (U - V)$

$\Rightarrow$ Kernal $= e^{-\gamma (U - V)^2} = \phi(\hat{U}) \phi(\hat{V})$

$= e^{-\gamma (U^2 + V^2 - 2UV)}$

$= e^{\gamma(-U^2 - V^2 + 2UV)}$

$= e^{\gamma(-U^2 - V^2)} \times e^{\gamma(2UV)}$

using Taylor expansion of $e^{\gamma(2UV)}$

$= \left(1 + \frac{\gamma 2UV}{1!} + \frac{(\gamma 2UV)^2}{2!} + \ldots\right)$

$= 1.1 + \frac{\sqrt{2\gamma} U . \sqrt{2\gamma} V}{1! \; 1!} + \sqrt{\frac{(2\gamma)^2}{2!}} U^2 \times \frac{\sqrt{2\gamma} . V^2}{\sqrt{2!}} + \ldots)$

where each term is a dot product of terms one from U & one from V.

$\Rightarrow K(U,V) = e^{-U^2\gamma} \times e^{-V^2\gamma} \times \left(1.1 + \frac{\sqrt{2\gamma} U}{1!} . \frac{\sqrt{2\gamma} V}{1!} + \ldots\right)$

$= e^{-U^2\gamma}\left(1, \frac{\sqrt{2\gamma} U}{1!}, \sqrt{\frac{(2\gamma)^2 U^2}{2!}} \ldots\right) . \left(e^{-V^2\gamma}\left(1, \frac{\sqrt{2\gamma} V}{1!} \ldots\right.\right.$

Scanned by CamScanner

Date
Page

$$\Rightarrow \phi(U) = e^{-\gamma U^2}\left(1, \sqrt{\frac{2\gamma}{1!}}U, \sqrt{\frac{(2\gamma)^2}{2!}}U^2 \ldots\right)$$

$$\phi(V) = e^{-\gamma V^2}\left(1, \sqrt{\frac{2\gamma}{1!}}V, \sqrt{\frac{(2\gamma)^2}{2!}}V^2 \ldots\right)$$

$$\Rightarrow K(U,V) = \phi(U).\phi(V)$$

$\Rightarrow$ Hence transform

$$\phi(x) = e^{-\gamma x^2}\left(1, \sqrt{\frac{2\gamma}{1!}}x, \sqrt{\frac{(2\gamma)^2}{2!}}x^2, \ldots\right)$$

as there is a factorial in the denominator, when the ~~deg~~ degree of the terms $\simeq$ ~~100000~~ 1000 000, then the denominator is going to be $\simeq \infty$, but the numerator is although quite big, it is not as big as the denominator. $\Rightarrow$ The terms goes to zero.

Hence higher degrees are close to zero (0) in rbf kernel.